

39302S1

CLAIMS

1. A method for link-level flow control, comprising:
 - establishing a plurality of logical links over a physical link between a transmitting entity and a receiving entity in a network;
 - assigning to the logical links respective maximum limits of transmission credits, the credits corresponding to space available to the links in a dynamically allocable portion of a receive buffer at the receiving entity, such that a sum of the maximum limits for all of the logical links corresponds to an amount of space substantially larger than a total volume of the space in the dynamically allocable portion of the receive buffer;
 - responsive to traffic from the transmitting entity to the receiving entity on a given one of the logical links, allocating one or more of the credits to the given logical link when it is determined that a total of the credits allocated to the given logical link is no greater than the respective maximum limit, and that a total of the credits allocated to all of the logical links together corresponds to an allocated volume that is no greater than the total volume of the space in the dynamically allocable portion of the receive buffer; and
 - controlling transmission of data over the given logical link responsive to the allocated credits.
2. A method according to claim 1, and comprising allocating to the logical links respective static portions of the receive buffer so as to hold, in each of the static portions, only the data received over the respective link, wherein the dynamically allocable portion of the receive buffer comprises the portion of

the buffer remaining after allocation of the static portions.

3. A method according to claim 2, wherein allocating the static portions comprises, when it is determined that a total of the credits allocated to the given logical link is greater than the respective maximum limit, or that a total of the credits allocated to all of the logical links together corresponds to an allocated volume that is greater than the total volume of the space in the dynamically allocable portion of the receive buffer, allocating the credits to the given logical link corresponding to the space available in the respective static portion of the receive buffer.

4. A method according to claim 3, wherein controlling the transmission of data comprises withholding transmission if the credits allocated are not sufficient to cover an amount of the data that is to be transmitted.

5. A method according to claim 3, and comprising:
receiving the data in the receive buffer responsive to the allocated credits;

passing the data from the receive buffer for onward transmission through the network; and

redistributing the allocated credits, after releasing the data, among the logical links having receive queues in the receive buffer that use both the static and dynamically allocable portions of the buffer.

6. A method according to claim 1, wherein controlling the transmission of data comprises withholding transmission if the credits allocated are not sufficient to cover an amount of the data that is to be transmitted.

7. A method according to claim 1, and comprising:
 - receiving the data in the receive buffer responsive to the allocated credits;
 - passing the data from the receive buffer for onward transmission through the network; and
 - releasing the allocated credits for use in subsequent transmission of data.
8. A method according to claim 7, wherein releasing the allocated credits comprises redistributing the released credits among the plurality of logical links.
9. A method according to claim 8, wherein redistributing the released credits comprises distributing the released credits in a round robin among at least some of the logical links.
10. A method according to claim 7, wherein allocating the one or more of the credits comprises allocating the credits from a pool of credits available to the logical links, and wherein releasing the allocated credits comprises returning the credits to the pool.
11. A method according to claim 1, wherein the network comprises a switching fabric, and wherein establishing the plurality of logical links comprises establishing virtual lanes in the switching fabric.
12. A method according to claim 11, wherein the switching fabric operates in accordance with an InfiniBand specification.
13. A method according to claim 1, wherein controlling the transmission of data comprises controlling the transmission of data packets by the transmitting entity.

14. A method according to claim 13, wherein controlling the transmission of the data packets comprises sending a flow control packet from the receiving entity to the transmitting entity, informing the transmitting entity of the allocated credits.

15. Network communication apparatus, comprising:
a transmitter, adapted to transmit data; and
a receiver, coupled to the transmitter by a physical link in the network and adapted to receive the data from the transmitter over a plurality of logical links established by the transmitter and the receiver over the physical link therebetween,

the receiver comprising a receive buffer with respect to which respective maximum limits of transmission credits are assigned to the logical links, the credits corresponding to space available to the links in a dynamically allocable portion of the receive buffer, such that a sum of the maximum limits for all of the logical links corresponds to an amount of space substantially larger than a total volume of the space in the dynamically allocable portion of the receive buffer,

the receiver being adapted, responsive to traffic from the transmitting entity to the receiving entity on a given one of the logical links, to allocate one or more of the credits to the given logical link when the receiver determines that a total of the credits allocated to the given logical link is no greater than the respective maximum limit, and that a total of the credits allocated to all of the logical links together corresponds to an allocated volume that is no greater than the total volume of the space in the dynamically allocable portion of the receive buffer, and

the transmitter being adapted to control transmission of the data over the given logical link responsive to the allocated credits.

16. Apparatus according to claim 15, wherein respective static portions of the receive buffer are allocated to the logical links so as to hold, in each of the static portions, only the data received over the respective link, wherein the dynamically allocable portion of the receive buffer comprises the portion of the buffer remaining after allocation of the static portions.

17. Apparatus according to claim 16, wherein the receiver is adapted to allocate the credits to the given logical link corresponding to the space available in the respective static portion of the receive buffer when the receiver determines that a total of the credits allocated to the given logical link is greater than the respective maximum limit, or that a total of the credits allocated to all of the logical links together corresponds to an allocated volume that is greater than the total volume of the space in the dynamically allocable portion of the receive buffer.

18. Apparatus according to claim 15, wherein the transmitter is adapted to withhold transmission of the data if the credits allocated are not sufficient to cover an amount of the data that is to be transmitted.

19. Apparatus according to claim 15, wherein the receiver is adapted to receive the data in the receive buffer responsive to the allocated credits, to pass the data from the receive buffer for onward transmission through the network, and to release the allocated credits

39302S1

after passing the data, for use in subsequent transmission of data.

20. Apparatus according to claim 19, wherein the receiver is adapted to redistribute the released credits among the plurality of logical links.

21. Apparatus according to claim 19, wherein the receiver is adapted to maintain a pool of credits available to the links, from which it allocates the credits to the logical links, and to return the released credits to the pool.

22. Apparatus according to claim 15, wherein the network comprises a switching fabric, and wherein the logical links comprise virtual lanes in the switching fabric.

23. Apparatus according to claim 22, wherein the switching fabric operates in accordance with an InfiniBand specification.

24. Apparatus according to claim 15, wherein the data comprise data packets.

25. Apparatus according to claim 24, wherein the receiver is adapted to send flow control packets to the transmitter, informing the transmitter of the allocated credits.